Hewlett - Packard Company
Legal Department, M/S 35
Intellectual Property Administration
P. O. Box 272400
Fort Collins, CO 80527-2400

Appl. No.	:	10/807,949	Confirmation No. 4075
Appellant	:	Tong Zhang	
Filed	:	March 24, 2004	
TC/A.U.	:	2600/2624	
Examiner	:	Sathyanaraya V Perungavoor	
Docket No.	:	200300641-1	
Customer No.		022879	
Title	:	Intelligent Key-Frame Extraction from a Video	

Mail Stop Appeal Briefs - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPELLANT'S BRIEF ON APPEAL

Sir:

This is an appeal under 37 CFR 41.31 from a Final Rejection in an Office Action mailed June 16, 2009. A Notice of Appeal was filed on September 15, 2009. An oral hearing was not requested.

Respectfully submitted,
TONG ZHANG

By: /J. Michael Johnson/
J. Michael Johnson, Reg. No. 37,856
Agent for Appellant(s)
Tel. (775) 849-3085/ Fax: (775) 849-3143
Email: jmjohnson@nsapatents.com

CERTIFICATE OF TRANSMISSION

I hereby certify that this correspondence is being transmitted to the United States Patent and Trademark Office via EFS-Web on the date shown below.

/J. Michael Johnson/ November 5, 2009
J. Michael Johnson Date

STATEMENT OF THE REAL PARTY IN INTEREST

The real party in interest is HEWLETT-PACKARD DEVELOPMENT COMPANY, of Houston, Texas, USA.

STATEMENT OF RELATED CASES

There are no known related prior or pending appeals, interferences or judicial proceedings.

JURISDICTIONAL STATEMENT

Jurisdiction over this appeal resides in the Board of Patent Appeals and Interferences under 35 U.S.C. §134(a). The Examiner issued a final rejection on June 16, 2009, setting a three month shortened statutory period for response. The time for responding to the final rejection expired on September 16, 2009. A Notice of Appeal was filed on September 15, 2009. The time for filing an appeal brief is two months after the filing of a notice of appeal. Bd. R. 41.37(c). The Appeal Brief is being filed on November 5, 2009.

TABLE OF CONTENTS

TABLE OF AUTHORITIES

A7. MPEP §2142, *Establishing s Prima Facie Case of Obviousness* ..... page 27.

A8. *KSR Int'l Co. v. Teleflex Inc.*, 127 S.Ct. 1727, 1741 (2007) ..... pages 27, 30 and 34.

A9. *In re Kahn*, 441 F.3d 977, 988, 78 USPQ2d 1329, 1336 (Fed. Cir. 2006) ..... pages 30 and 34.

A10. *In re McLaughlin*, 443 F.2d 1392, 170 USPQ 209 (CCPA 1971) ..... page 31.

A11. *In re Royka,* 490 F.2d 981, 985 (CCPA 1974) ..... pages 31 and 34.

STATUS OF AMENDMENTS

No amendments, whether entered or not, have been filed after final rejection.

GROUNDS OF REJECTION TO BE REVIEWED

For each ground of rejection that Appellant contests hereinbelow, wherein the rejection ground applies to more than one claim, the additional claims do not stand or fall together to the extent the claims are separately identified and argued below.

Ground 1: Rejection of Claims 23, 25, 26, 28-33, 36, 38, 39 and 41-46 under 35 U.S.C. 102(e) as being anticipated by Toklu et al., U.S. Patent No. 6,549,643 B1 (hereinafter 'Toklu et al.').

Ground 2: Rejection of Claims 24, 27, 34, 37, 40 and 47 under 35 U.S.C. 103(a) as being unpatentable over Toklu et al. in view of Wu et al., U.S Patent Application Publication No. 2003/0068087 A1 (hereinafter 'Wu et al.').

Ground 3: Rejection of Claims 35 and 48 under 35 U.S.C. 103(a) as being unpatentable over Toklu et al. in view of Dimitrova et al., U.S. Patent No. 6,125,229 (hereinafter 'Dimitrova et al.').

5

STATEMENT OF FACTS

Facts with respect to Ground 1

(1) The Examiner contended that Toklu et al., anticipated a key-frame extraction system as recited in Appellant's Claim 23. See Final Office Action (hereinafter 'FOA'), item [3], page 4-5. In particular, the Examiner contended that Toklu et al. disclose a Video Segmenter 12 (Toklu et al., FIG. 1) that "extracts each of a series of video frames …," identifying the Video Segmenter as being equivalent to a "video frame extractor" and further equating "segments" defined by segment boundaries produced by the Video Segmenter 12 to a "series of video frames [extracted by the video frame extractor] from a video," as recited in Claim 23 (FOA, item [3], page 4). Referring to Toklu et al., the Examiner further contended that operating levels of a Key-Frame Selection module 13 (including ref. nos. 15, 16 and 17) illustrated in FIG. 1 were equivalent to "a set of frame analyzers that obtain the series of video frames in parallel …, each frame analyzer selecting a corresponding set of candidate key-frames from the series…," as recited in Claim 23 (FOA, item [3], page 4). Finally, the Examiner contended that a Key Frame Eliminator 18 disclosed by Toklu et al. was equivalent to a "key-frame selector that … arranges the candidate key-frames into a set of clusters and that selects one of the candidate key-frames from each cluster as a key-frame for the video," as further recited in Claim 23. See FOA, item [3], page 4, last line to

page 5, lines 1-4. While Appellant agrees that an output of the Video Segmenter 12 according to Toklu et al. may be video frames arranged in segments, Appellant otherwise disagrees with the Examiner.

Toklu et al. do not disclose "a set of frame analyzers" that operate in parallel to perform "different corresponding analysis" where each frame analyzer selects "a corresponding set of candidate key-frames," as recited in Claim 23. Instead, Toklu et al. disclose in FIG. 1 a Process Selection 14 module that determines (e.g., see Toklu et al., FIG. 2A, step 202, "Select Appropriate Operating Level…") which of four identified operating levels (designated 'A,' 'B,' 'C' and 'D') of the Key-Frame Selection module 13 are to be performed at any one time by the Key-Frame Selection module 13 to produce a single set of key-frames for a segment or unit (see Toklu et al., step 210, FIG. 2A, "Output Key-Frame List for Each Unit"). Only level D involves more than one type of analyses (i.e., both "Motion Activity" and "Color Histogram Activity" analyses) and the two types are explicitly disclosed as being performed in series. While Toklu et al. allow for "one or more of the key-frame selection processes 15, 16, 17 to be selected …," Toklu et al. only disclose one configuration (i.e., level D) that includes more than one of the processes. Moreover, while Toklu et al. make vague reference to the two processes of level D being "performed one at a time (in any order) or in parallel," Toklu et al. never actually disclose a configuration in which the two processes (i.e.,

15 and 17) are performed in parallel. Regardless, even if performed in parallel, the result of operating level D as well as the results of the other operating levels (i.e., A, B and C) disclosed by Toklu et al. are a single list of key-frames at step 210. As such, Toklu et al. never actually disclose a set of key frame analyzers operating in parallel and certainly fail to disclose, "each frame analyzer selecting a corresponding set of candidate key-frames," as claimed by Appellant.

Further, Toklu et al. fail to disclose a "key-frame selector" that either or both "arranges the candidate key-frames into a set of clusters," and "selects one of the candidate key-frames from each cluster as a key-frame for the video," as recited in Claim 23. In particular, Toklu et al. are silent on arranging key-frames into clusters and further fail to disclose selecting only ***one*** key-frame from each cluster.

Instead, according to Toklu et al., the Key-Frame Elimination module 18 processes key-frames, "to eliminate key-frames that are visually similar" (Col. 6, lines 32-33). "After the key-frames have been selected (steps 207 or 208) for each segment, a key-frame elimination process [e.g., performed by the Key-Frame Elimination module 18] … is performed (step 209) to eliminate the selected key-frames that are visually similar" (Toklu et al., Col. 7, lines 56-59). Toklu et al. disclose no other function for the Key-Frame Elimination module 18 beyond eliminating visually similar key-frames. Specifically, Toklu et al. never disclose, implicitly or explicitly, that the Key-Frame Elimination module

18 either "arranges the candidate key-frames into a set of clusters" or "selects one of the candidate key-frames from each cluster as a key frame for the video," contrary to the Examiner's contention. Further, Toklu et al. never disclose, implicitly or explicitly, such "arranging" or "selecting" is or may be performed after key-frame selection by any of the types of analyses that are identified (e.g., Figs. 2A-2C of Toklu et al.).

In fact, Toklu et al. do not disclose or even suggest clusters of key-frames. Toklu et al. do disclose "video segments" or "shots" defined by segment boundaries which, "represent a contiguous recording of one or more video frames depicting a continuous action in time and space," (Toklu et al., Col. 5, lines 55-57). However, there is never a reference by Toklu et al. of any particular grouping of key-frames other than a natural grouping associated with the processing of segments.

The Examiner contended that "segments" according to Toklu et al. were equivalent to "clusters," as recited in Appellant's claims. However, according to Toklu et al., segments are exclusively produced by the Video Segmenter 12 and are only "arranged" in as much as they might be related to an overall sequential structure of the video from which the segments are produced. Toklu et al. never mention arranging the segments and certainly fail to disclose a module that

arranges the segments, including but not limited to the Key-Frame Elimination module 18 relied upon by the Examiner.

Moreover, the video segments are provided along with the input video data (as segment boundaries) to the Key-Frame Selection module 13, according to Toklu et al. See for example, Col. 5, lines 61-65. As such, the segments according to Toklu et al. are created before key-frames are selected. Thus, the segments of Toklu et al. are not and respectfully cannot be equivalent to "clusters" produced by arranging candidate key frames, as claimed by Appellant.

Finally, the according to Toklu et al., the Key-Frame Elimination module 18 produces a list of ***one or more*** key-frames per processed segment. For example, according to disclosed embodiments of a key-frame elimination process illustrated in FIG. 3, step 305 and FIG. 4, step 407, the process comprises selecting "the Reference Image as a Key-Frame," and designating "the Comparison Key-Frame as The New Reference Key-Frame," for further comparison with other key-frames in the segment. As such, unless there happened to be only one key-frame that survived the elimination process (an unlikely event), the key-frame elimination processes of FIGs. 3 and 4 will always yield a plurality of key-frames. In particular, these processes are not specifically limited to producing only one key-frame per segment in clear contrast to Appellant's recitation in Claim 23 of selecting "one of the candidate key-frames from each cluster".

(2) Regarding Claim 25, the Examiner contended that Toklu et al. disclosed a "key-frame selector (*i.e., 18*)" that "selects the key-frames by determining an importance score (*i.e., object motion*) for each candidate key-frame in each cluster (*i.e., segment*)," further citing FIG. 4 and Col. 13, lines 60-66 (FOA, page 5, second paragraph). Appellant disagrees.

The term, 'object motion' as employed by Toklu et al. is not equivalent to an importance score, according to Appellant's definition and usage thereof. In particular, an 'importance score' is defined as a score "based on a set of characteristics of the candidate key-frame," that may include among a number of different characteristics "whether the candidate key-frame satisfies one of the camera motion rules of the camera motion tracker" (Appellant's specification, page 21, line 25 to page 23, line 12). On the other hand, while Toklu et al. acknowledge that a good key-frame selection method may "exploit the dynamic information contained in videos due to camera and/or object motion," (Toklu et al., Col. 3, lines 17-22), Toklu et al. do not actually employ object motion in key-frame selection. For example, key-frame selection may employ a computed motion activity (MA) curve that is "generated using the estimated [***camera***] motion parameters (step 213)" (Toklu et al., Col. 9, lines 20-21 and FIG. 2B, ***emphasis*** added). At Col. 13, lines 60-66, Toklu et al. disclose, as part of key-frame elimination, using a morphologically smoothed difference image to "find the overall object motion

(step 404)" and making a decision regarding whether or not the object motion is significant. As such, Toklu et al. disclose making a choice between two alternatives (significant motion or insignificant motion), which is clearly different than selecting based on an importance score, as claimed by Appellant. In fact, Toklu et al. are silent on using an importance score for key-frame selection, relying instead on other criteria for selection and elimination of key-frames.

(3) Regarding Claim 26, the Examiner contended that Toklu et al. disclosed "importance scores (*i.e., object motion)*" determined by "determining an image content (i.e., region segmentation) of each candidate key-frame," citing FIG. 4 and Col 13, lines 37-40. Appellant disagrees.

As noted above, Toklu et al. do not disclose the use of object motion as an importance score. Further, at Col. 13, lines 37-40, Toklu et al. disclose selecting a "reference image (preferably the first key-frame in the ordered list)" and further disclose segmenting "the image into regions of similar color (step 401)." Segmenting the image into regions of similar color is not equivalent to "determining an image content" which generally includes much more than just regions of similar color.

(4) Regarding Claim 28, the Examiner contended that Toklu et al. disclosed selecting key frames "by determining an image quality (*i.e., object motion*) … ." Appellant disagrees.

Appellant defined "image quality" used in determining an "image quality score" as being "based on the sharpness of the candidate key-frame or on the brightness of the candidate key-frame or a combination of sharpness and brightness" (Appellant's specification, page 23, lines 17-20). As such, "image quality" used by the key-frame selector to select key-frames clearly is not equivalent to "object motion," as the Examiner contended. Toklu et al. do not disclose the use of image quality for key-frame selection or elimination.

(5) Regarding Claim 30, the Examiner contended that Toklu et al. disclosed a key-frame extraction system "wherein the frame analyzers include a color layout (*i.e., distribution or histogram*) analyzer," referencing FIG. 1, ref. no. 17. Appellant disagrees.

A color layout analyzer as recited in Claim 30 analyzes a relative location of different colors (color layout) within an image. For example, the color layout may be defined in terms of an average color in each of a plurality of sub-blocks distributed across the image (e.g., the image divided into an NxN grid of sub-blocks). See discussion beginning at page 11, line 7 through page 13, line 3, of Appellant's specification. On the other hand, a 'color histogram' as employed by both Toklu et al. and Appellant is defined (according to conventional usage in the art) as a representation of a frequency distribution of colors in an image. A color histogram is generally produced or derived by counting, for each color in a color

space of the image, a number of image pixels having that color. For example, there may be 25 pixels having a first color (e.g., blue), 102 pixels having a second color (e.g., yellow), and so on. Together, the exemplary pixel counts comprise a color histogram. An exemplary color histogram may be displayed as a bar graph where each bar of a plurality of bars represents a particular color and a length of each bar is determined according to the respective pixel counts for each color. Clearly a "color layout" as recited in Claim 30 is not equivalent to the color histogram disclosed by Toklu et al. Further, Toklu et al. do not disclose a color layout as claimed by Appellant.

(6) Regarding Claim 31, the Examiner contended Toklu et al. disclosed "a fast camera motion detector," citing FIG. 1, ref. no. 15. Appellant disagrees.

According to Appellant's specification at page 13, lines 15-26, "[t]he fast camera motion detector may detect a fast motion of the camera that captured the video 12 by detecting a relatively large difference in the color layouts or the color histograms of adjacent video frames over a number of consecutive video frames … video frames in the video 12 that correspond to periods of fast camera motion are not selected because fast motion tends to blur images." Toklu et al. disclose that the Motion Detection and Analysis module 15 employ estimated camera motion parameters (e.g., see FIG. 2B, step 213) and examines whether motion detected exceeds various thresholds including thresholds associated with panning, tilting,

absolute zooming (e.g., see FIG. 2B, steps 220, 221, 222). However, Toklu et al. do not specifically disclose detecting "fast camera motion".

(7) Regarding Claim 33, the Examiner contended Toklu et al. disclosed "an object motion analyzer," referring to FIG. 1, ref. no. 15. Appellant disagrees.

According to Toklu et al., the Motion Detection and Analysis module 15 only detects and analyzes camera motion. The key-frame elimination process (i.e., Key Frame Elimination module 18) does employ motion of the object. However, the object motion is only employed in terms of whether or not the object motion is considered significant (e.g., see Toklu et al., FIG. 4, step 405). Toklu et al. fail to disclose an object motion analyzer as defined in Appellant's specification, page 16, line 14 to page 18, line 12. For example, Toklu et al. do not disclose anything that "examines the trajectories of moving objects in the video ...," (Appellant's specification page 16, lines 14-17) or that "tracks it [the moving object] through the video frames of the video 12 ..." (Appellant's specification page 18, lines 7-9).

(8) Regarding Claims 36, 38, 39 and 41-46, the Examiner relied upon discussion with respect to Claims 23, 25, 26 and 28-33 to contend that Toklu et al. disclosed all claim limitations. Appellant disagrees.

In particular, Toklu et al. do not disclose "selecting multiple sets of candidate key-frames from a video including detecting ***multiple types of meaningful content*** in the video by performing in ***parallel a set of different***

analyses on each video frame in the video," "arranging the candidate key-frames into a set of clusters," and "selecting ***one*** of the candidate key-frames from each cluster …," as recited in Claim 36 (***emphasis*** added). Further, Toklu et al. fail to disclose using an "importance score" for "selecting one of the candidate key-frames from each cluster," as recited in Claim 38 or "determining an importance score" comprising "determining image content …," as recited in Claim 39. Similarly, Toklu et al. fail to disclose all of the elements recited in at least Claim 41 (e.g., "determining an image quality"), Claim 42 (e.g., "performing in ***parallel*** a set of different analyses," including "performing a color histogram analysis," ***emphasis*** added), Claim 43 (e.g., "performing in parallel a set of different analyses includes performing a color layout analysis"), Claim 44 (e.g., "performing in parallel a set of different analyses includes performing a fast camera motion analysis"), Claim 45 (e.g., "performing in ***parallel*** a set of different analyses includes performing a camera motion detection," ***emphasis*** added), and Claim 46 (e.g., "performing in parallel a set of different analyses includes performing an object motion track").

Facts with respect to Ground 2

(9) With respect to Claims 24, 27, 34, 37, 40 and 47, the Examiner admitted that Toklu et al. fail to explicitly disclose "an audio event detector …," as variously recited in Claims 24 and 37; or "determining an audio content of each

candidate key-frame," as recited in Claims 27 and 40; or a "human face detector," as variously recited in Claims 34 and 47. The Examiner contended that Wu et al. disclose elements and limitations (e.g., audio analysis and face detection) recited in the subject claims that are admittedly explicitly lacking in Toklu et al. The Examiner concluded that it would have been obvious to modify the teachings of Toklu et al. with Wu et al. to include audio analysis and face detection. Further, the Examiner contended a skilled artisan would have found "motivation" to include audio analysis according to Wu et al. because it would "reduce the amount of image processing performed on the video data" (FOA, page 8, lines 6 and 20). Similarly, the Examiner contended that inclusion of face detection according to Wu et al. would be motivated because "human faces are most important users of video content" (FOA, page 9, line 11). Appellant disagrees.

Regarding the audio analysis, Wu et al. disclose an audio analyzing algorithm 501 that facilitates distinguishing audio data fragments that include human voices from audio data fragments that include "non-human sounds, such as noises or silence," such that corresponding "image data fragments can be separated" and not processed using a disclosed face-detection-analyzing algorithm (Wu et al, paragraph [0027]). However, Wu et al. do not disclose an audio event detector as defined and claimed by Appellant. For example, Wu et al. do not disclose an audio event detector that "detects audio events … that may indicate a

highlight," such as, "applause, screaming, acclaim, [and] the start of high level noise after a period of silence" (Appellant's specification, page 18, lines 30-32). Wu et al. disclose only the detection of "human sounds."

Further, regarding Claim 24, Wu et al. fail to disclose providing a set of selected candidate key-frames to a key-frame selector. Instead, Wu et al. explicitly disclose that audio-analyzing algorithm 501 is used to decrease an amount of data that is to be processed by a face-detection-analyzing algorithm 503. In particular, the only identified recipient of any output the audio-analyzing algorithm 501 is the aforementioned face-detection-analyzing algorithm 503.

Regarding Claim 27, Wu et al. do not disclose determining an importance score "by determining an audio content," as recited therein. Regarding using the audio-analyzing algorithm 501, Wu et al. only disclose deciding whether or not to further process frames using the face-detection-analyzing algorithm 503 based on detection of human sounds (or the lack thereof) in the audio data fragment associated with a particular image data fragment (see Wu et al., paragraph [0027]).

Facts with respect to Ground 3

(10) With respect to Claims 35 and 48, the Examiner admitted that Toklu et al. fail to explicitly disclose "a user interface for displaying a set of video frames in the video previous to … and … subsequent to each key-frame and for obtaining a user selection of one or more the video frames," as is variously recited therein.

The Examiner contended that Dimitrova et al. disclosed the elements lacking in Toklu et al., citing Col. 12, lines 59-67, and concluded it would be obvious to modify Toklu et al. with Dimitrova et al. because "the user's desire should be taken into account" (FOA, page 10, line 11). Appellant disagrees.

At Col. 12, lines 59-67, Dimitrova et al. disclose that a source video may be "stopped once the video index has been read," that "keyframes" may be "displayed on a display," that "the visual index may be printed if the user desires," and that "[a] user may also select a particular keyframe from the visual index," (i.e., step 816). Further, Dimitrova et al. disclose, "[i]f a user wishes to view the source video at that particular keyframe, the source tape could be automatically forwarded to a corresponding point on the source tape from where the keyframe was extracted …" (Dimitrova et al, Col. 12, line 65 to Col. 13, line 2). However, Dimitrova et al. at least never mention displaying a set of video frames "previous to each key-frame" and other than allowing a user to play the source tape starting at a keyframe, Dimitrova et al. are silent on "selecting one or more of the video frames," as is variously recited in Claims 35 and 48.

In as much as certain facts associated with respect to Grounds 1-3 are not explicitly discussed above, Appellant believes that these facts are more properly a subject of an argument and, as such, are presented below in an *Arguments* section.

ARGUMENTS

Ground 1:

Claims 23, 25, 26, 28-33, 36, 38, 39 and 41-46: (Argument previously presented in a Response/Amendment filed 04/02/09) The Examiner erred in finding that Appellant's Claims 23, 25, 26, 28-33, 36, 38, 39 and 41-46 are anticipated by Toklu et al. under 35 U.S.C. 102(e). In particular, the Examiner failed to present and support a *prima facie* case of anticipation with respect to Toklu et al.

In order to establish *prima facie* anticipation and maintain a rejection based thereon, a single prior art reference must describe, either expressly or inherently, each and every element as set forth in the claim. *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987). "The identical invention must be shown in as complete detail as is contained in the … claim." *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989). "The elements must be arranged as required by the claim, but this is not an *ipsissimis verbis* test, i.e., identity of terminology is not required. *In re Bond*, 910 F.2d 831, 15 USPQ2d 1566 (Fed. Cir. 1990)" MPEP §2131 *Anticipation – Application of 35 U.S.C. 102(a), (b), and (e) [R-1]*. Further, the Examiner is obliged to consider all words in a claim in judging the patentability of that claim against the prior art. *In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970).

It is respectfully submitted that the Examiner either failed to consider all of the elements and limitations of Appellant's rejected claims or has mischaracterized or misunderstood aspects of the disclosure of Toklu et al. As such, the Examiner failed to establish that each and every element recited in the rejected claims is, in fact, disclosed by Toklu et al. Further, Appellant submits that the differences between the Toklu et al. reference and Appellant's claimed invention would be clear to a skilled artisan and are not merely differences in terminology.

For example, regarding independent base Claim 23, Toklu et al. at least fail to disclose a key frame selector "that obtains the corresponding candidate key-frames from each frame analyzer and arranges the candidate key-frames into a set of clusters," and further, "that selects one of the candidate key-frames from each cluster as a key-frame for the video," as is recited therein. Toklu et al. further do not disclose a set of frame analyzers "that obtain the series of video frames in parallel from the video extractor," as recited in Claim 23. Regarding independent base Claim 36, Toklu et al. at least do not disclose, "arranging the candidate key-frames into a set of clusters," and, "selecting one of the candidate key-frames from each cluster as a key-frame for the video," as recited therein. Toklu et al. also fail to disclose, "selecting multiple sets of candidate key-frames," and do not explicitly enable, "performing in ***parallel*** a set of different analyses on each video frame in

the video" (***emphasis*** added). Merely mentioning that something may be done in parallel does not enable one to practice a parallel implementation.

In an *Examiner's Response* portion of a Response to Arguments section of the FOA, page 3, item 1.a., the Examiner contended that Toklu et al., at steps 400 and 409 of FIG 4, teach that "key frames grouped according to corresponding segments represent a cluster." However, as has been discussed above, segments according to Toklu et al. are exclusively produced by a Video Segmenter 12. As such, Toklu et al. do not and respectively cannot disclose, "a key-frame selector that obtains … and ***arranges*** the candidate key-frames into a set of clusters," as recited in Claim 23. In particular, the segments according to Toklu et al. are already arranged prior to any attempt to extract key-frames. Similarly, Toklu et al. do not disclose, "arranging the candidate key-frames into a set of clusters," since an arrangement of segments is dictated by their respective production according to an operation of the Video Segmenter 12.

In the *Examiner's Response*, the Examiner further contended that selecting a reference key frame "as the keyframe" at step 407 of FIG. 4 was equivalent to selecting "one of the candidate key-frames from each cluster as the key-frame for the video." However, Toklu et al. do not disclose selecting ***only*** one key-frame. In particular, at step 407 Toklu et al. selects a reference key-frame as a key-frame and then continues processing by designating a comparison key-frame as the new

reference key frame. Since the process according to Toklu et al. then continues with other key-frames from the key-frame set for a given segment, the likelihood is that multiple reference key-frames will be selected as a key-frame in step 407. Regardless, neither at step 407 nor anywhere else therein do Toklu et al. explicitly limit a number of selected key-frames for a given segment to just one.

Hence, there is no support for a finding of *prima facie* anticipation since Toklu et al. clearly fail to disclose each and every element as set forth and arranged in the base Claims 23 and 36. Claims 25, 26 and 28-33 are dependent from and include all of the limitations of base Claim 23. Claims 38, 39 and 41-46 are dependent from and include all of the limitations of base Claim 36. A finding of *prima facie* anticipation in view of Toklu et al. cannot be made for Appellant's dependent Claims 25, 25, 28-33, 38, 39 and 41-46 for at least the same reasons provided above for Appellant's base Claims 23 and 36, respectively.

Claims 25 and 38: (New Argument) The Examiner erred in finding that all of the elements and limitations recited exclusively in Appellant's Claims 25 and 38 are disclosed by Toklu et al. Specifically, Toklu et al. fail to disclose an importance score or the use thereof. Without disclosure of an importance score, Toklu et al. do not and respectfully cannot support a finding of *prima facie* anticipation of Appellant's Claims 25 and 38, when considered separately from the other claims.

Claims 26 and 39: (New Argument) The Examiner erred in finding that all of the elements and limitations recited exclusively in Appellant's Claims 26 and 39 are disclosed by Toklu et al. In particular, Toklu et al. fail to disclose an importance score that is determined by an image content of each candidate key-frame. As such, Toklu et al. do not and respectfully cannot support a finding of *prima facie* anticipation of Appellant's Claims 26 and 39, when considered separately from the other claims.

Claims 28 and 41: (New Argument) The Examiner erred in finding that all of the elements and limitations recited exclusively in Appellant's Claims 28 and 41 are disclosed by Toklu et al. In particular, Toklu et al. fail to disclose determining an image quality for each candidate key-frame in each cluster. In fact, Toklu et al. is silent on image quality. As such, Toklu et al. do not and respectfully cannot support a finding of *prima facie* anticipation of Appellant's Claims 28 and 41, when considered separately from the other claims.

Claims 30 and 43: (New Argument) The Examiner erred in finding that all of the elements and limitations recited exclusively in Appellant's Claims 30 and 43 are disclosed by Toklu et al. In particular, Toklu et al. fail to disclose a color layout analyzer or performing color layout analysis. Toklu et al. do not consider color layout. As such, Toklu et al. do not and respectfully cannot support a finding

of *prima facie* anticipation of Appellant's Claims 30 and 43, when considered separately from the other claims.

Claims 31 and 44: (New Argument) The Examiner erred in finding that all of the elements and limitations recited exclusively in Appellant's Claims 31 and 44 are disclosed by Toklu et al. In particular, Toklu et al. fail to disclose a fast camera motion detector or performing a fast camera motion analysis. As such, Toklu et al. do not and respectfully cannot support a finding of *prima facie* anticipation of Appellant's Claims 31 and 44, when considered separately from the other claims.

Claims 33 and 46: (New Argument) The Examiner erred in finding that all of the elements and limitations recited exclusively in Appellant's Claims 33 and 46 are disclosed by Toklu et al. In particular, Toklu et al. fail to disclose an object motion analyzer or performing an object motion track. Toklu et al. are silent on tracking a motion of an object. As such, Toklu et al. do not and respectfully cannot support a finding of *prima facie* anticipation of Appellant's Claims 33 and 46, when considered separately from the other claims.

Accordingly, Appellant respectfully requests that the Board of Patent Appeals and Interferences find for the Appellant and reverse the rejection of Appellant's Claims 23, 25, 26, 28-33, 36, 38, 39 and 41-46 under 35 U.S.C. 102(e) as being anticipated by Toklu et al.

Ground 2:

Claims 24, 27, 34 37, 40 and 47: (Argument previously presented in a Response/Amendment filed 04/02/09) The Examiner erred in performing the *Graham* factual inquires (*Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966)) and further that the Examiner has not presented and properly supported a finding of obviousness with respect to the rejected claims.

In particular, the Examiner provided no indication in the written record that the *Graham* factual inquires had been made and that the results thereof had been applied. For example, the Examiner at least failed to show the level of ordinary skill in the pertinent art and objective evidence relevant to the issue of obviousness. Without a level of ordinary skill being identified and supported by facts, it is not possible to properly provide objective evidence relevant to the issue of obviousness. Moreover, it is inappropriate to find "[i]t would have been obvious to one of ordinary skill in the art …," to do anything with respect to the teachings of one or more of the cited references without a level of ordinary skill having first been established.

In addition, the Examiner erred by failing to provide evidence, which is deemed critical by the courts, to establish that it would have been obvious to the skilled artisan at the time of the invention to combine these references in the ways that the Examiner proposed. In particular, the Examiner failed to clearly articulate

or make explicit the reason(s) why the invention, as claimed, would have been obvious. See MPEP §2142, *Establishing a Prima Facie Case of Obviousness* ("Analysis supporting a rejection under 35 U.S.C. 103 should be made explicit"). No such analysis is evident in the present Office Action.

In particular, an invention, "composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art." *KSR Int'l Co. v. Teleflex Inc.*, 127 S.Ct. 1727, 1741 (2007). There is neither teaching in the cited references nor explanation based on scientific reasoning that a skilled artisan would have found it obvious to combine teachings of Toklu et al. with those of Wu et al., as proposed. Further, there is no teaching, suggestion or motivation or even a discussion of any other factors that might have led the skilled artisan to combine the references, as required by the court. *Id.* As noted by the court, "it can be important to identify a reason that would have prompted a person of ordinary skill in the relevant field to combine the [prior art] elements in the way the claimed new invention does." *Id.* "To facilitate review, this analysis should be made explicit." *Id.* As such, the Examiner erred by not providing any evidence that a skilled artisan had "apparent reason to combine the known elements in the fashion claimed" (*Id.* at 1741) beyond merely identifying in the prior the alleged presence of various claimed elements and

making a general but unsupported statement regarding an alleged advantage (considered a motivation by the Examiner) of the combination.

The Examiner suggested that a "motivation" for modifying the teachings of Toklu et al. by including the audio analysis of Wu et al. would be, "to reduce the amount of image processing performed on the video data." However, according to Wu et al., "[t]he audio-analyzing algorithm 501 is used to analyze the audio data 411 of the video data 41 so that audio data fragments with human voice included in the audio data 411 and their corresponding image data fragment in the image data 412 are screened" (paragraph [0027]). By screening the video data for frames associated with human voice sounds, Wu et al. are able to reduce, "the amount of the video data processed by the face-detection-analyzing algorithm 503" (paragraph [0026]). In other words, only when human voices are detected is there a ***need*** to process the video data using the face-detection-analyzing algorithm. As such, reduced amount of data processing, according to Wu et al., is only associated with face detection and is merely a result of deciding whether a video frame is likely to have a human face to detect. "Therefore, the audio data fragments of non-human sounds, such as noises or silence, and their corresponding image data fragments can be separated, and no process using the face-detection-analyzing algorithm is performed." Wu et al., paragraph [0027].

As such, combining the audio analysis of Wu et al. ***would not*** actually, "reduce the amount of image processing performed on the video data," of Toklu et al., contrary to the Examiner's contention. In particular, the image processing reduction realized by Wu et al. is only associated with face detection which Toklu et al. do not disclose using. In fact, adding audio analysis according to Wu et al. to the teachings of Toklu et al. would necessarily increase, not decrease, the amount of processing, albeit perhaps not image processing. The Examiner's motivation regarding a potential reduction in image processing is respectfully without merit.

The Examiner's motivation for combining the face detection of Wu et al. with the teachings of Toklu et al. is likewise respectfully without merit. In particular, the Examiner's stated motivation that, "human faces are most important users of video content," respectfully does not make sense. It is not clear how a human face (or faces) are or even can be important users of video content. Even if the Examiner meant to say that human faces are the most important features to users of the video content, the motivation still respectfully lacks merit. Human faces are typically only ***one factor or feature*** among many that may be important in generating a key-frame from a video, a fact recognized by both Toklu et al. and Wu et al. Toklu et al. disclose motion detection, pixel-based frame detection, and color histogram computation, as ways of generating key-frames. Moreover, Toklu et al. did not even consider face detection, explicitly or implicitly.

Wu et al. similarly disclose other analyses besides face detection (e.g., the shot-shift-analyzing algorithm 502) that would seem to argue against human faces being ***most important*** in selecting key-frames. Moreover, there is effectively no support within either of the references for a conclusion that, in either Toklu et al. or Wu et al., the presence of a human face in a video frame was the ***most important*** feature (including, but not limited to Wu et al., paragraph [0005] cited by the Examiner).

In fact, the Examiner's rationale for combining the aforementioned references is essentially nothing more than a respectfully incoherent conclusory statement. "[R]ejections on obviousness cannot be sustained with mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *In re Kahn*, 441 F.3d 977, 988, 78 USPQ2d 1329, 1336 (Fed. Cir. 2006). See also *KSR Int'l Co. v. Teleflex Inc.*, cited *supra* (quoting Federal Circuit statement with approval). Furthermore, it is respectfully submitted that the Examiner appears to be simply using Appellant's invention as a template for a hindsight reconstruction of Appellant's claims. The court made clear, "[a] factfinder should be aware, of course, of the distortion caused by hindsight bias and must be cautious of argument reliant upon ex post reasoning." *KSR Int'l Co. v. Teleflex Inc.*, cited *supra*. Of course, some hindsight reconstruction ***may*** be unavoidable. However, just because

some hindsight reconstruction may be unavoidable, the Court's acknowledgment of the fact does not provide the Examiner the liberty to use such hindsight. Instead, the Examiner is obliged to demonstrate with reasoned explanation and solid evidence that hindsight reconstruction in view of Applicant's disclosure was not the sole basis for the Examiner's conclusion regarding the obviousness to combine the references. *In re McLaughlin*, 443 F.2d 1392, 170 USPQ 209 (CCPA 1971).

Notwithstanding the above, the combination of Toklu et al. and Wu et al. further fail to disclose or suggest all of the claim limitations, contrary to the Examiner's contention. *In re Royka,* 490 F.2d 981, 985 (CCPA 1974). In particular, Toklu et al. at least fail to disclose or suggest all of the limitations of base Claim 23 and separately, base Claim 36, as are discussed above. Wu et al. fail to add the elements that are lacking in Toklu et al. with respect to the limitations of Applicant's base claims. Claims 24, 27 and 34 are dependent from and include all of the limitations of base Claim 23. Claims 37, 40 and 47 are dependent from and include all of the limitations of amended base Claim 36. Therefore, the combination of Toklu et al. and Wu et al. fails to disclose or suggest all of the limitations of either the base Claims 23 and 36 or dependent Claims 24, 27, 34, 37, 40 and 47. Further, a combination of Toklu et al. and Wu et al. fails to disclose or suggest all of the limitations exclusively found in each of dependent

Claims 24, 27, 34, 37, 40 and 47, respectively. A detailed discussion of the missing elements and limitations is provided above in the *STATEMENT OF FACTS* section and are not repeated here for brevity only. A lack of disclosure or suggestion of claim limitations by the combined references still further defeats the *prima facie* case of obviousness with respect to the dependent claims.

Hence, the rejection of Claims 24, 27, 34, 37, 40 and 47 under 35 U.S.C. 103(a) as being unpatentable over Toklu et al. in view of Wu et al. lacks the elements necessary for proper support of a *prima facie* case of obviousness. Accordingly, Appellant respectfully requests that the Board of Patent Appeals and Interferences find for the Appellant and reverse the rejection of Appellant's Claims 24, 27, 34, 37, 40 and 47.

Ground 3:

Claims 35 and 48: (Argument previously presented in a Response/Amendment filed 04/02/09) The Examiner erred in finding that Appellant's Claims 35 and 48 are unpatentable under 35 U.S.C. 103(a) over Toklu et al. in view of Dimitrova et al. In particular, the rejection lacks support for a finding of obviousness under 35 U.S.C. 103(a) including, but not limited to, explicit evidence that the *Graham* factual inquires (*Graham v. John Deere Co.*, cited *supra*) were performed and that the results were properly applied.

In particular, the Examiner provided no indication in the written record that the *Graham* factual inquires had been made and the results thereof have been applied. For example, Applicant submits that the Examiner at least failed to show the level of ordinary skill in the pertinent art and objective evidence relevant to the issue of obviousness. Without a level of ordinary skill being identified and supported by facts, it is not possible for the Examiner to properly provide objective evidence relevant to the issue of obviousness.

Furthermore, nothing either in the references themselves or in that provided by the Examiner establishes how the system and method for selecting key-frames of video data disclosed by Toklu et al. is or could be modified to include or employ the various elements extracted by the Examiner from the disclosure of Dimitrova et al. Nothing in the record establishes how Toklu et al. and Dimitrova et al. may be combined to render Applicant's rejected claims *prima facie* obvious as proposed by the Examiner. Without evidence in the record that such a substitution or modification is even possible and that there is a reasonable expectation of success in the combination that would lead to the result relied upon by the Examiner, the proposed modification of the disclosure of Toklu et al. using elements of Dimitrova et al. lacks proper support for a finding of *prima facie* obviousness.

Finally, the Examiner's rationale for combining the aforementioned references is essentially nothing more than a conclusory statement. Other than

acknowledging a desirability of taking into account a "user's desire," the Examiner has not articulated any reasoning for the combination beyond identifying in Dimitrova et al. an element or limitation that is admittedly explicitly lacking in Toklu et al. when considering Applicant's claims. As noted above, "rejections on obviousness cannot be sustained with mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *In re Kahn*, cited *supra*. In fact, it is respectfully submitted that the Examiner appears to be using Applicant's invention as a template for a hindsight reconstruction of Applicant's claims. See *KSR Int'l Co. v. Teleflex Inc.*, cited *supra*. Hence, there is insufficient support in the written record for a finding by the Examiner of *prima facie* obviousness.

Notwithstanding the above, the combination of Toklu et al. and Dimitrova et al. further fail to disclose or suggest all of the claim limitations, contrary to the Examiner's contention. *In re Royka,* 490 F.2d 981, 985 (CCPA 1974) In particular, Toklu et al. at least fail to disclose or suggest all of the limitations of base Claim 23 and separately, base Claim 36, as is discussed above. Dimitrova et al. fail to add the elements that are lacking in Toklu et al. with respect to the limitations of Applicant's base claims. Claim 35 is dependent from and includes all of the limitations of base Claim 23. Claim 48 is dependent from and includes all of the limitations of base Claim 36. Therefore, the combination of Toklu et al.

and Dimitrova et al. fails to disclose or suggest all of the limitations of either the base Claims 23 and 36 or the dependent Claims 35 and 48. Further, a combination of Toklu et al. and Dimitrova et al. fails to disclose or suggest all of the limitations exclusively found in each of dependent Claims 35 and 48, respectively. A detailed discussion of the missing elements and limitations of Claims 35 and 48 is provided above in the *STATEMENT OF FACTS* section and is not repeated here for sake of brevity only. A lack of disclosure or suggestion of claim limitations by the combined references still further defeats the *prima facie* case of obviousness with respect to the dependent claims.

Hence, the rejection of Claims 35 and 48 under 35 U.S.C. 103(a) as being unpatentable over Toklu et al. in view of Dimitrova et al. lacks the elements necessary for proper support of a *prima facie* case of obviousness. Accordingly, Appellant respectfully requests that the Board of Patent Appeals and Interferences find for the Appellant and reverse the rejection of Appellant's Claims 35 and 48 under 35 U.S.C. 103(a).

APPENDIX

Claims:

Claim 23 (Rejected): A key-frame extraction system, comprising:

video frame extractor that extracts each of a series of video frames from a video;

a set of frame analyzers that obtain the series of video frames in parallel from the video frame extractor, each frame analyzer selecting a corresponding set of candidate key-frames from the series by performing a different corresponding analysis on each video frame in the series such that the analyses are selected to detect multiple types of meaningful content in the video;

key-frame selector that obtains the corresponding candidate key-frames from each frame analyzer and arranges the candidate key-frames into a set of clusters and that selects one of the candidate key-frames from each cluster as a key-frame for the video.

Claim 24 (Rejected): The key-frame extraction system of claim 23, further comprising an audio event detector that obtains the series of video frames from the video frame extractor and that selects a corresponding set of candidate key-frames from the series by performing an audio analysis on each video frame in the series and that provides the corresponding set of candidate key-frames to the key-frame selector.

Claim 25 (Rejected): The key-frame extraction system of claim 23, wherein the key-frame selector selects the key-frames by determining an importance score for each candidate key-frame in each cluster.

Claim 26 (Rejected): The key-frame extraction system of claim 25, wherein the key-frame selector determines the importance scores by determining an image content of each candidate key-frame.

Claim 27 (Rejected): The key-frame extraction system of claim 25, wherein the key-frame selector determines the importance scores by determining an audio content of each candidate key-frame.

Claim 28 (Rejected): The key-frame extraction system of claim 23, wherein the key-frame selector selects the key-frames by determining an image quality for each candidate key-frame in each cluster.

Claim 29 (Rejected): The key-frame extraction system of claim 23, wherein the frame analyzers include a color histogram analyzer.

Claim 30 (Rejected): The key-frame extraction system of claim 23, wherein the frame analyzers include a color layout analyzer.

Claim 31 (Rejected): The key-frame extraction system of claim 23, wherein the frame analyzers include a fast camera motion detector.

Claim 32 (Rejected): The key-frame extraction system of claim 23, wherein the frame analyzers include a camera motion tracker.

Claim 33 (Rejected): The key-frame extraction system of claim 23, wherein the frame analyzers include an object motion analyzer.

Claim 34 (Rejected): The key-frame extraction system of claim 23, wherein the frame analyzers include a human face detector.

Claim 35 (Rejected): The key-frame extraction system of claim 23, further comprising a user interface for displaying a set of video frames in the video previous to each key-frame and a set of video frames in the video subsequent to each key-frame and for obtaining a user selection of one or more of the video frames.

Claim 36 (Rejected): A method for key-frame extraction, comprising:

selecting multiple sets of candidate key-frames from a video including detecting multiple types of meaningful content in the video by performing in parallel a set of different analyses on each video frame in the video;

arranging the candidate key-frames into a set of clusters; and

selecting one of the candidate key-frames from each cluster as a key-frame for the video,

wherein arranging the candidate key-frames and selecting one of the candidate key-frames are performed by a key-frame extraction system, and wherein the key-frame extraction system outputs the selected one key-frame from each cluster as the key-frames for the video.

Claim 37 (Rejected): The method of claim 36, wherein selecting multiple sets of candidate key-frames includes selecting a set of candidate key-frames from the video by performing an audio analysis on each video frame in the video.

Claim 38 (Rejected): The method of claim 36, wherein selecting one of the candidate key-frames from each cluster includes determining an importance score for each candidate key-frame in each cluster.

Claim 39 (Rejected): The method of claim 38, wherein determining an importance score comprises determining an image content of each candidate key-frame.

Claim 40 (Rejected): The method of claim 38, wherein determining an importance score comprises determining an audio content of each candidate key-frame.

Claim 41 (Rejected): The method of claim 36, wherein selecting one of the candidate key-frames from each cluster includes determining an image quality for each candidate key-frame in each cluster.

Claim 42 (Rejected): The method of claim 36, wherein performing in parallel a set of different analyses includes performing a color histogram analysis.

Claim 43 (Rejected): The method of claim 36, wherein performing in parallel a set of different analyses includes performing a color layout analysis.

Claim 44 (Rejected): The method of claim 36, wherein performing in parallel a set of different analyses includes performing a fast camera motion analysis.

Claim 45 (Rejected): The method of claim 36, wherein performing in parallel a set of different analyses includes performing a camera motion detection.

Claim 46 (Rejected): The method of claim 36, wherein performing in parallel a set of different analyses includes performing an object motion track.

Claim 47 (Rejected): The method of claim 36, wherein performing in parallel a set of different analyses includes performing a human face detection.

Claim 48 (Rejected): The method of claim 36, further comprising displaying a set of video frames in the video previous to each key-frame and a set of video frames in the video subsequent to each key-frame and obtaining a user selection of one or more of the video frames.

5

Claims Support and Drawing Analysis:

Claim 23: A key-frame extraction system **{FIG 2, 10; Page 6, line 29 to Page 7, line 27; Page 24, lines 10-26}**, comprising:

video frame extractor **{FIG 2, 14; Page 7, lines 1-13}** that extracts each of a series of video frames from a video **{FIG 2, 12; Page 7, lines 1-13}**;

a set of frame analyzers **{FIG 2, 20-24; Page 7, lines 1-13, lines 29-30; Page 11, lines 7-9; Page 13, lines 14-15 and 28-29; Page 16, lines 13-14, Page 18, lines 14-15}** that obtain the series of video frames in parallel from the video frame extractor **{FIG 2, 14; Page 7, lines 1-12}**, each frame analyzer **{FIG 2, 20-24}** selecting a corresponding set of candidate key-frames **{FIG 2, 18; Page 7, lines 12, 18 and 20-27; Page 19, line 2; Page 19, line 28 to Page 20, line 26}** from the series by performing a different corresponding analysis on each video frame in the series such that the analyses are selected to detect multiple types of meaningful content in the video **{FIG 2, 12}**;

key-frame selector **{FIG 2, 30; Page 7, lines 20-27; Page 20, line 7 to Page 24, line 26}** that obtains the corresponding candidate key-frames **{FIG 2, 18}** from each frame analyzer **{FIG 2, 20-24}** and arranges the candidate key-frames **{FIG 2, 18}** into a set of clusters **{Page 6, lines 16-19}** and that selects one of the candidate key-frames **{FIG 2, 18}** from each cluster as a key-frame for the video **{FIG 2, 12}**.

Claim 24: The key-frame extraction system **{FIG 2, 10}** of claim 23, further comprising an audio event detector **{FIG 2, 16; Page 7, lines 14-18; Page 18, line 28 to Page 19, line 26; Page 23, lines 5-8}** that obtains the series of video frames from the video frame extractor **{FIG 2, 14; Page 7, lines 1-13}** and that selects a corresponding set of candidate key-frames **{FIG 2, 18}** from the series by performing an audio analysis **{Page 19, line 8 to Page 20, line 7; FIGs. 8a-8b}** on each video frame in the series and that provides the corresponding set of candidate key-frames to the key-frame selector **{FIG 2, 30; Page 7, lines 20-27; Page 20, line 7 to Page 24, lines 26}**.

Claim 25: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the key-frame selector **{FIG 2, 30}** selects the key-frames by determining an importance score **{Page 21, line 23 to Page 24, line 26, FIG 9, 202}** for each candidate key-frame **{FIG 2, 18}** in each cluster.

Claim 26: The key-frame extraction system **{FIG 2, 10}** of claim 25, wherein the key-frame selector **{FIG 2, 30}** determines the importance scores by determining an image content of each candidate key-frame **{FIG 2, 18}**.

Claim 27: The key-frame extraction system **{FIG 2, 10}** of claim 25, wherein the key-frame selector **{FIG 2, 30}** determines the importance scores by determining an audio content of each candidate key-frame **{FIG 2, 18}**.

Claim 28: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the key-frame selector **{FIG 2, 30}** selects the key-frames by determining an image quality **{FIG 9, 204}** for each candidate key-frame **{FIG 2, 18}** in each cluster.

Claim 29: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the frame analyzers **{FIG 2, 20-24}** include a color histogram analyzer **{Page 7, line 29 to Page 11, line 5; FIG. 3}**.

Claim 30: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the frame analyzers **{FIG 2, 20-24}** include a color layout analyzer **{Page 11, line 7 to Page 13, line 12; FIG. 4}**.

Claim 31: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the frame analyzers **{FIG 2, 20-24}** include a fast camera motion detector **{Page 13, lines 14-26}**.

Claim 32: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the frame analyzers **{FIG 2, 20-24}** include a camera motion tracker **{Page 13, line 28 to Page 16, line 11; FIGs. 5a-5c}**.

Claim 33: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the frame analyzers **{FIG 2, 20-24}** include an object motion analyzer **{Page 16, line 13 to Page 18, line 12; FIGs. 6, 7a-7b}**.

Claim 34: The key-frame extraction system **{FIG 2, 10}** of claim 23, wherein the frame analyzers **{FIG 2, 20-24}** include a human face detector **{Page 18, lines 14-26}**.

Claim 35: The key-frame extraction system **{FIG 2, 10}** of claim 23, further comprising a user interface **{Page 24, lines 21-26}** for displaying a set of video frames in the video **{FIG 2, 12}** previous to each key-frame and a set of video frames in the video **{FIG 2, 12}** subsequent to each key-frame and for obtaining a user selection of one or more of the video frames.

Claim 36: A method **{FIG. 1, Page 6, lines 3-27}** for key-frame extraction, comprising:

selecting **{FIG. 1, 300, Page 6, lines 5-14}** multiple sets of candidate key-frames **{FIG 2, 18}** from a video **{FIG 2, 12}** including detecting multiple types of meaningful content in the video **{FIG 2, 12; Page 7, lines 1-13}** by performing in parallel a set of different analyses on each video frame in the video **{FIG 2, 12}**;

arranging **{FIG. 1, 302, Page 6, lines 16-19}** the candidate key-frames **{FIG 2, 18}** into a set of clusters **{Page 6, lines 16-19}**; and

selecting **{FIG. 1, 304, Page 6, lines 21-27}** one of the candidate key-frames **{FIG 2, 18}** from each cluster as a key-frame for the video **{FIG 2, 12}**,

wherein arranging the candidate key-frames **{FIG 2, 18}** and selecting one of the candidate key-frames **{FIG 2, 18}** are performed by a key-frame extraction

system **{FIG 2, 10; Page 6, line 29 to Page 7, line 27; Page 24, lines 10-26}**, and wherein the key-frame extraction system **{FIG 2, 10}** outputs the selected one key-frame from each cluster as the key-frames for the video **{FIG 2, 12}**.

Claim 37: The method **{FIG. 1}** of claim 36, wherein selecting **{FIG. 1, 300}** multiple sets of candidate key-frames **{FIG 2, 18}** includes selecting a set of candidate key-frames **{FIG 2, 18}** from the video **{FIG 2, 12}** by performing an audio analysis **{Page 18, line 28 to Page 19, line 26; FIGs. 8a-8b}** on each video frame in the video **{FIG 2, 12}**.

Claim 38: The method **{FIG. 1}** of claim 36, wherein selecting one of the candidate key-frames **{FIG 2, 18}** from each cluster includes determining an importance score **{FIG 9, 202; Page 21, line 23 to Page 24, line 26}** for each candidate key-frame **{FIG 2, 18}** in each cluster.

Claim 39: The method **{FIG. 1}** of claim 38, wherein determining an importance score **{FIG 9, 202; Page 21, line 23 to Page 24, line 26}** comprises determining an image content of each candidate key-frame **{FIG 2, 18}**.

Claim 40: The method **{FIG. 1}** of claim 38, wherein determining an importance score **{FIG 9, 202; Page 21, line 23 to Page 24, line 26}** comprises determining an audio content of each candidate key-frame **{FIG 2, 18}**.

Claim 41: The method **{FIG. 1}** of claim 36, wherein selecting one of the candidate key-frames **{FIG 2, 18}** from each cluster includes determining an image quality **{FIG 9, 204; Page 23, lines 14 to Page 24, line 8}** for each candidate key-frame **{FIG 2, 18}** in each cluster.

Claim 42: The method **{FIG. 1}** of claim 36, wherein performing in parallel a set of different analyses includes performing a color histogram analysis **{Page 7, line 29 to Page 11, line 5; FIG. 3}**.

Claim 43: The method **{FIG. 1}** of claim 36, wherein performing in parallel a set of different analyses includes performing a color layout analysis **{Page 11, line 7 to Page 13, line 12; FIG. 4}**.

Claim 44: The method **{FIG. 1}** of claim 36, wherein performing in parallel a set of different analyses includes performing a fast camera motion analysis **{Page 13, lines 14-26}**.

Claim 45: The method **{FIG. 1}** of claim 36, wherein performing in parallel a set of different analyses includes performing a camera motion detection **{Page 13, line 28 to Page 16, line 11; FIGs. 5a-5c}**.

Claim 46: The method **{FIG. 1}** of claim 36, wherein performing in parallel a set of different analyses includes performing an object motion track **{Page 16, line 13 to Page 18, line 12; FIGs. 7a-7b}**.

Claim 47: The method **{FIG. 1}** of claim 36, wherein performing in parallel a set of different analyses includes performing a human face detection **{Page 18, lines 14-26}**.

Claim 48: The method **{FIG. 1}** of claim 36, further comprising displaying a set of video frames in the video **{FIG 2, 12}** previous to each key-frame and a set of video frames in the video **{FIG 2, 12}** subsequent to each key-frame and obtaining a user selection of one or more of the video frames.

Means or step plus function:

There are no claims that employ means-plus-function construction.

Evidence Appendix:

None.

Related Proceedings Appendix:

None.

* * * * *